

# MENU SEARCH INDEX DETAIL

1/1



# PATENT ABSTRACTS OF JAPAN

(11)Publication number: 06274394

(43)Date of publication of application: 30.09.1994

(51)Int.CI.

G06F 12/00 G06F 13/00

(21)Application number: 05059963

(71)Applicant:

NIPPON TELEGR & TELEPH CORP <NTT>
FUJIMURA TAKASHI

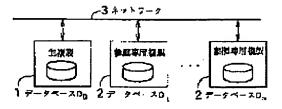
(22)Date of filing: 19.03.1993

(72)Inventor:

(54) DISTRIBUTED DATA BASE CONTROL METHOD

## (57)Abstract:

PURPOSE: To quickly and securely transfer update information to plural reference dedicated data base by transferring update information to be transferred by appending a circular list storing identifiers of all the reference dedicated data bases. CONSTITUTION: A data base 1 which is permitted to update and is provided with a main duplicate and N-pieces of data bases provided with reference dedicated duplicated can communicate data between optional data bases through a network 3. Then, update information to be transferred is added with the circular list storing identifiers of all the reference dedicated data base 2 and while circulating update information appended with the circular list among all the reference data bases 2, update information is transferred to all the reference data bases 2. Consequently, even when including the one which can not communicate because of a fault, etc., the other party in a communication possible state can dynamically be selected from the circular list to quickly inform in order from the communication possible one.



LEGAL STATUS

# (書誌+要約+請求の範囲)

```
(19)
     【発行国】日本国特許庁(JP)
    【公報種別】特許公報(B2)
【特許番号】第2740105号
【登録日】平成10年(1998)1月23日
(12)
(11)
(24)
(45)
     【発行日】平成10年(1998)4月15日
    【発明の名称】分散デー
                           -タベース制御方法
(54)
(51) 【国際特許分類第6版】
   G06F 12/00
        13/00
                  355
[F | ]
   G06F 12/00
                  533 J
        13/00
                  355
 【請求項の数】3
 【全頁数】 6
(21) 【出願番号】特願平5-59963
(22) 【出願日】平成5年(1993)3月19日
(65) 【公開番号】特開平6-274394
(43) 【公開日】平成6年(1994)9月30日
【審查請求日】平成7年(1995)1月30日
(73) 【特許権者】
【識別番号】 000004226
 【氏名又は名称】日本電信電話株式会社
 【住所又は居所】東京都新宿区西新宿三丁目19番2号
(72) 【発明者】
【氏名】藤村 考
 【住所又は居所】東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
(74) 【代理人】
 【弁理士】
 【氏名又は名称】三好 秀和 (外1名)
 【審査官】 田川 泰宏
(56)【参考文献】
 [文献] 特開 平4-302338 (JP, A)
【文献】特開 平4-302040 (JP, A)
【文献】特開 平4-52736 (JP, A)
【文献】特開 平3-103942 (JP, A)
【文献】特開 平5-35560 (JP, A)
【文献】特開 平4-318637 (JP, A)
```

# (19)日本国特許庁(JP)

# (12) 公開特許公報(A)

FΙ

(11)特許出願公開番号

特開平6-274394

(43)公開日 平成6年(1994)9月30日

(51)Int.Cl.5

識別記号

庁内整理番号

技術表示箇所

G 0 6 F 12/00

5 3 3 J 8944-5B

13/00

3 5 5

7368-5B

審査請求 未請求 請求項の数3 OL (全 6 頁)

(21)出願番号

(22)出願日

特願平5-59963

平成5年(1993)3月19日

(71)出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(72) 発明者 藤村 考

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

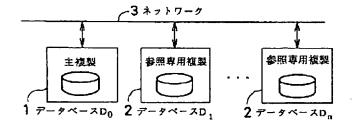
(74)代理人 弁理士 三好 秀和 (外1名)

# (54) 【発明の名称 】 分散データベース制御方法

# (57) 【要約】

【目的】 ネットワークにおけるトラヒックの集中やネ ットワークの部分的な障害による更新情報の伝達の遅れ を低減しながら、複数の参照専用データベースに更新情 報を迅速かつ適確に転送し得る分散データベース制御方 法を提供する。

【構成】 転送すべき更新情報にすべての参照専用デー タベースの識別子を格納した回覧リストを付与し、この 回覧リストを付与された更新情報をすべての参照専用デ ータベースの間で回覧しながら、すべての参照専用デー タベースに更新情報を転送する。



1

## 【特許請求の範囲】

【請求項1】 更新を許されている主複製データを有する主複製データベースと前記主複製データを複製したデータを参照専用として保持する複数の参照専用データベースとが通信ネットワークによって互いに接続された分散データベースにおいて前記主複製データベースに対して行われた更新情報を前記複数の参照専用データベースのすべてに転送するように制御する分散データベース制御方法であって、各参照専用データベースに転送すべき更新情報に前記複数の参照専用データベースのすべての識別子を格納した回覧リストを付与し、この回覧リストを付与された前記更新情報を前記複数の参照専用データベースの間で回覧して、すべての参照専用データベースに更新情報を転送することを特徴とする分散データベース制御方法。

【請求項2】 前記更新情報の緊急性に応じて更新情報の前記回覧リストを分割し、この分割された回覧リストに従って複数の参照専用データベース間に同時に更新情報を転送することを特徴とする請求項1記載の分散データベース制御方法。

【請求項3】 前記通信ネットワークの負荷に応じて更新情報の前記回覧リストを分割し、この分割された回覧リストに従って複数の参照専用データベース間に同時に更新情報を転送することを特徴とする請求項1記載の分散データベース制御方法。

#### 【発明の詳細な説明】

# [0001]

【産業上の利用分野】本発明は、更新を許されている主複製データベースに対して行われた更新情報を複数の参照専用データベースのすべてに効率的かつ適確に転送する分散データベース制御方法に関し、特に更新情報の緊急性やネットワークの負荷に応じて複数の参照専用データベースに対する更新情報の転送を適切に制御する分散データベース制御方法に関する。

#### [0002]

【従来の技術】電話番号案内業務やクレジット・カードの信用情報確認処理のように、データベースに対する参照要求が更新要求に比較して極めて高い頻度で発生する業務では、データベースの内容の最新性を犠牲にしても参照処理の高い並列性が必要となる場合がある。このよりな業務では、従来から参照専用のデータベースの複製を複数個作り、それを地域毎等の単位で分散配置して利用可能性を高めるとともに、更新に関してはデータの一貫性を保つために主複製という一つのデータベースのみに制限するという技術が用いられてきた。つまり、更新は主複製に対してのみ許し、更新があったときはその更新情報をその他の主複製でないすべての複製に転送し更新するというものである。この場合、主複製に対する更新がその他の複製に反映されるまでの間は、古い複製が参照される可能性があるが、これは参照処理の利用可能50

性を高めるための犠牲とみなし許容している。このような方式を用いているシステムには、例えば、「Tichy, W. F., and Ruan, Z.1984. Towardsa distributed file system. In Proceedings of Usenix 1984 Summer Conference, pp. 87-97」にある Ibisファイルシステムをはじめとしていくつか存在する。

【0003】このような方式のデータベースでは更新情報をこれらの複製に対して効率的に通知することが重要であり、従来からいくつかの方法が使われている。例えば、主複製に対して更新が行なわれる度に、他の参照専用の複製にその更新情報を送る方式では、ネットワークのトラヒックが膨大になり許容できない場合があるため、主複製に対する更新が行なわれる度ではなく、一定時間、一定量あるいはオペレータの指定などの契機で、ある程度の量の更新情報をまとめて参照専用の複製に転送し、一括して更新処理を行なう方法が使われている。【0004】

【発明が解決しようとする課題】しかしながら、上述した従来の方法において、参照専用の複製の数が極めて多い場合、例えば、数百から数千以上の複製がある場合は、主複製を持つデータベースがすべての複製に対して、更新情報を同報(broadcast )する方法では、同報を行ったときにネットワークの負荷が急速に高まるため、ネットワークの負荷が安定しないという問題と、主複製を持つデータベースにこの処理の負荷が集中するという問題があるが、この問題も図5のように複製を持つデータベースを階層化し中間層のデータベースが更新情報の転送を分担するような従来方式によって、ある程度解決することが可能である。

0 【0005】しかし、このような方法では中間層のデータベースが何らかの障害により、停止しているときに、その下位のすべてのデータベースに更新情報が伝わらないという問題がある。

【0006】本発明は、上記に鑑みてなされたもので、その目的とするところは、ネットワークにおけるトラヒックの集中やネットワークの部分的な障害による更新情報の伝達の遅れを低減しながら、複数の参照専用データベースに更新情報を迅速かつ適確に転送し得る分散データベース制御方法を提供することにある。

#### 0 [0007]

【課題を解決するための手段】上記目的を達成するため、本発明の分散データベース制御方法は、更新を許されている主複製データを有する主複製データベースと前記主複製データを複製したデータを参照専用として保持する複数の参照専用データベースとが通信ネットワークによって互いに接続された分散データベースにおいて前記主複製データベースに対して行われた更新情報を前記複数の参照専用データベース間御方法であって、各参照専用データベースに転送すべき更新情報に前記複数の参照

3

専用データベースのすべての識別子を格納した回覧リストを付与し、この回覧リストを付与された前記更新情報を前記複数の参照専用データベースの間で回覧して、すべての参照専用データベースに更新情報を転送することを要旨とする。

# [0008]

【作用】本発明の分散データベース制御方法では、転送すべき更新情報にすべての参照専用データベースの識別子を格納した回覧リストを付与し、この回覧リストを付与された更新情報をすべての参照専用データベースの間 10で回覧し、すべての参照専用データベースに更新情報を転送する。

#### [0009]

【実施例】以下、図面を用いて本発明の実施例を説明する。

【0010】図1は、本発明の一実施例に係わる分散データベース制御方法が適用される分散データベースの構成を示すブロック図である。同図において、1は更新を許されている主複製を持つデータベース、2は参照専用複製を持つN個のデータベース、3はこれらの任意のデータベース間でデータの通信を可能にするネットワークである。オペレータやアプリケーションからのデータベースに対する更新要求およびロックが必要な参照要求はすべて主複製を持つデータベースに対して送られ処理される。また、それ以外の参照要求はN個の参照専用データベースの内の一つに送られて処理される。N個の参照専用データベースの内の一つに送られて処理される。N個の参照専用データベースの中のどれを選択するかということは、例えば、参照専用データベースを地域毎に配置した場合は、その地域内で発生した参照要求をそれぞれのデータベースに割り当てるという方法がある。

【0011】図2は主複製を持つデータベースの機能を 示す論理ブロック図であり、4はデータの主複製、5は 参照/更新要求を一時的に蓄積する入力キュー、6は参 照/更新要求の回答を蓄積する出力キュー、7は更新さ れたレコードを一時的に蓄積する更新レコードキュー、 8 は参照専用データベースに回覧する更新情報を一時的 に蓄積する更新情報出力キュー、9は参照/更新要求を 入力キュー5から取り出し処理し、その結果を出力キュ 一6につなぎ、もしこの処理が主複製の更新を伴うとき は、更新レコード(更新履歴レコード)を更新レコード 40 キュー7につなぐ参照/更新処理部、10はネットワー クの負荷状態を監視するネットワーク負荷監視部、11 は更新レコードキューから一つあるいは複数個の更新レ コードを取り出し、図3に示すデータ構造の一つあるい は複数個の更新情報13を作り、更新情報出力キュー8 につなぐ回覧処理部、12はネットワーク3から参照/ 更新要求を受け取り、入力キュー5につなぎ、その回答 を出力キュー6から取り出し送信する機能と、更新情報 出力キュー8から更新情報を取り出し、参照専用データ ベースに送信する機能を持つ通信処理部である。

【0012】参照/更新処理部9は、従来技術のデータベース管理システム(DBMS)で実現でき、ネットワーク負荷監視部10も従来技術で実現できるので詳細は省略する。

【 0 0 1 3】回覧処理部 1 1 は、以下のステップで実現できる。

【0014】ステップS11:更新レコードキュー7に つながれている更新レコードがある一定量に到達する か、あるいは一定時間経過する度に、以下を繰り返す。

【0015】ステップS12:すべての緊急度(1...K)について、以下を繰り返す(ただし、緊急度はK段階に分類されていると仮定する)。

【0016】ステップS13:更新レコードキュー7につながれている更新レコードの中からその緊急度を有するものを集めた更新レコードリストを作る。また、この更新レコードリストに対して更新情報連続番号を割り当てる。更新情報連続番号は割り当てられる度に1づつ増加し、同じ番号が割り当てられることはない。

【0017】ステップS14:参照専用のすべてのデータベースの識別子から成る図3の13で示すような回覧リストを作る。

【0018】ステップS15:緊急度Xとネットワークの負荷状態Yから回覧リストの分割数Dを計算する。Dは例えばINT(C\*(X/K)\*(L/Y))+1で計算できる。ただし、Cは定数、YはL段階に分類され負荷が大きい程高い値となるものと仮定する。これにより、ネットワークの負荷が低ければ低いほど、また緊急度が高ければ高い程、大きなDを得ることができる。

【0019】ステップS16:ステップS14で得られた回覧リストをステップS15で得られた分割数Dの数だけ分割する。割り切れないときは余りの回覧レコードを均等に分割された回覧リストに振り分ける。分割されたすべての回覧リストについて、ステップS13で得られた更新レコードリストと、更新情報連続番号と、緊急度を付与し、図3に示す構造を持つD個の更新情報の構造体を作り、それらを更新情報出力キュー8にストアする。

【0020】通信処理部12は、参照/更新要求の入出 力処理とともに、次のステップで更新情報の送信を行 う。

【0021】ステップS21:更新情報出力キュー8に 更新情報がつながれている間、以下を繰り返す。

【0022】ステップS22: 更新情報出力キュー8から先頭の更新情報を取り出し、その更新情報の回覧リストの中に通信可能状態にあるデータベースが存在するときは、その中の一つを選択し更新情報を送信する。もし、回覧リストの中に通信可能状態にあるデータベースが存在しないときは、更新情報出力キュー8の末尾にその更新情報を戻す。

50 【0023】図4は参照専用の複製を持つデータベース

の機能を示す論理ブロック図であり、18は参照専用デ ータの複製、19は参照要求を一時的に蓄積する入力キ ュー、20は参照要求の回答を蓄積する出力キュー、2 1は受信した更新情報を一時的に蓄積する更新情報入力 キュー、22は他の参照専用データベースに回覧する更 新情報を一時的に蓄積する更新情報出力キュー、23は 参照要求を入力キュー19から取り出し処理しその結果 を出力キュー20につなぐ参照処理部、24は更新情報 入力キュー21から更新情報を取り出し、更新処理を行 う更新処理部、25はネットワークの負荷状態を監視す 10 るネットワーク負荷監視部、26は更新情報入力キュー 21から更新情報を取り出し、その更新情報を加工し、 更新情報出力キュー22につなぐ回覧処理部、27はネ ットワーク3から参照要求を受け取り、入力キュー19 につなぎ、その回答を出力キュー20から取り出し送信 する機能と、更新情報出力キュー22から更新情報を取 り出し、他の参照専用データベースに送信する機能を持 つ通信処理部である。

【0024】参照処理部23は従来技術のデータベース 管理システム(DBMS)で実現でき、ネットワーク負 20 荷監視部25も従来技術で実現できるので詳細は省略す る。

【0025】更新処理部24は図3に示す構造を持つ更 新情報を受け取り、その更新情報に基づいて参照専用複 製の更新処理を行う。ただし、更新情報は回覧の経路や 緊急度の違いなどの理由によって必ずしも古いものから 順に届くわけではないことに注意が必要である。古い更 新情報S:に含まれるある更新レコードR: と、新しい 更新情報S: に含まれる更新レコードR: が同一のデー タレコードD(データベースの更新の単位)に対するも のであった場合に、R1 がR2 よりも先行して実行され るか、あるいはR2 のみが行われるように制御する必要 がある。この制御は、参照専用複製のすべてのデータレ コードに、そのデータレコードを更新した最新の更新レ コードの更新情報連続番号を格納する領域を付与し、デ ータレコードDを更新レコードRで更新を行うときに、 Dに付与されている更新情報連続番号No と、Rの更新 情報連続番号N®を比較し、N® <N® のときのみ、更 新することで実現できる。

【0026】回覧処理部26は以下のステップで実行する。

【0027】ステップS31:更新情報入力キュー22 につながれている更新情報13がある間、以下を繰り返 す。

【0028】ステップS32:更新情報入力キュー22 から更新情報13を取り出す。

【0029】ステップS33: 更新情報の回覧リストから自分のデータベースの識別子を取り除く。

【0030】ステップS34:更新情報の緊急度Xとネ べき更新情報に参照専用のすべてのデータベースの識別ットワークの負荷状態Yから回覧リストの分割数Dを計 50 子を格納した回覧リストを作り、更新情報の緊急性に応

算する。Dは例えばINT(C・(X/K)・(L/Y))+1で計算できる。ただし、Cは定数、YはL段階に分類され負荷が大きい程高い値となるものと仮定する。これにより、ネットワークの負荷が低ければ低いほど、また緊急度が高ければ高い程、大きなDを得ることができる。

【0031】ステップS35:更新情報の回覧リストをステップS34で得られた分割数Dの数だけ分割し、D個の回覧リストを作る。割り切れないときは余りの回覧レコードを均等に分割された回覧リストに振り分ける。分割されたそれぞれの回覧リストに、更新レコードリストと、更新情報連続番号と、緊急度を付与し、図3に示す構造を持つD個の更新情報の構造体を作り、それらを更新情報出力キュー22にストアする。

【0032】尚、更新処理部24と回覧処理部26はいずれも更新情報入力キュー21から更新情報を取り出しているが、この取り出しには別々のポインタを用い、更新処理部24と回覧処理部のいずれも同じ更新情報を受け取ることとする。

【0033】通信処理部27は、基本的に主複製を持つデータベースの通信処理部12と同様である。ただし、主複製を持つデータベースあるいは他の参照専用複製を持つデータベースから更新情報が送られた場合、それを更新情報入力キュー21につなぐ処理を追加する。

【0034】次に、以上のように構成される分散データベース制御方式の全体の処理を概略的に説明する。

【0035】主複製に対して行なわれた更新情報を他の 参照専用の複製を保持する複数のデータベースに転送す る場合においては、主複製を持つデータベースは、その 転送すべき更新情報に参照専用のすべてのデータベース の識別子を格納した回覧リストを作り、回覧リストの中 から通信可能状態にあるデータベースを一つ選択し、そ のデータベースにその更新情報と回覧リストを転送し、 回覧リストの中に通信可能状態にあるデータベースが存 在しないときは、一定時間経過後に再びそのデータベー スに対して更新情報と回覧リストの転送を試みる。

【0036】更新情報とそれに付与されている回覧リストを受け取った参照専用のデータベースは、回覧リストから自分のデータベースの識別子を取り除き、回覧リストの中から通信可能状態にあるデータベースを一つ選択し、そのデータベースにその更新情報と回覧リストを転送し、回覧リストの中に通信可能状態にあるデータベースがないときは、一定時間経過後に再びそのデータベースに対してその更新情報と回覧リストの転送を試みる。

【0037】主複製に対して行なわれた更新情報を他の 参照専用の複製を保持する複数のデータベースに転送す る場合において、特に、緊急性の高い更新情報を転送す る場合には、主複製を持つデータベースは、その転送す べき更新情報に参照専用のすべてのデータベースの識別 子を終始した同覧リストを作り、更新情報の緊急性に応

じて細かく更新情報の回覧リストを分割し、それぞれの 回覧リストの中から通信可能状態にあるデータベースを 一つ選択し、それらのデータベースにその更新情報と分 割された回覧リストをそれぞれ転送し、回覧リストの中 に通信可能状態にあるデータベースが存在しないとき は、一定時間経過後に再びそれらのデータベースに対し て更新情報と分割された回覧リストの転送を試みる。

【0038】更新情報とそれに付与されている回覧リス トを受け取った参照専用のデータベースは、回覧リスト から自分のデータベースの識別子を取り除き、その回覧 10 リストを更新情報の緊急性に応じて細かく分割し、それ ぞれの回覧リストに対して、通信可能状態にあるデータ ベースが存在するときは、その中から一つ選択し、その データベースにその更新情報と分割された回覧リストを 転送し、通信可能状態にあるデータベースがないとき は、一定時間経過後に再びそのデータベースに対してそ の更新情報と分割された回覧リストの転送を試みる。

【0039】主複製に対して行なわれた更新情報を他の 参照専用の複製を保持する複数のデータベースに転送す る場合において、特に、ネットワークの負荷が低いとき には、主複製を持つデータベースは、その転送すべき更 新情報に参照専用のすべてのデータベースの識別子を格 納した回覧リストを作り、ネットワークの負荷に応じて ネットワークの負荷の低いとき程細かく更新情報の回覧 リストを分割し、それぞれの回覧リストの中から通信可 能状態にあるデータベースを一つ選択し、それらのデー タベースにその更新情報と分割された回覧リストをそれ ぞれ転送し、回覧リストの中に通信可能状態にあるデー タベースが存在しないときは、一定時間経過後に再びそ れらのデータベースに対して更新情報と分割された回覧 30 リストの転送を試みる。

【0040】更新情報とそれに付与されている回覧リス トを受け取った参照専用のデータベースは、回覧リスト から自分のデータベースの識別子を取り除き、その回覧 リストをネットワークの負荷に応じてネットワークの負 荷の低いとき程細かく分割し、それぞれの回覧リストに 対して、通信可能状態にあるデータベースが存在すると きは、その中から一つ選択し、そのデータベースにその 更新情報と分割された回覧リストを転送し、通信可能状 態にあるデータベースがないときは、一定時間経過後に 再びそのデータベースに対してその更新情報と分割され た回覧リストの転送を試みる。

#### [0041]

【発明の効果】以上説明したように、本発明によれば、 転送すべき更新情報にすべての参照専用データベースの 識別子を格納した回覧リストを付与し、この回覧リスト を付与された更新情報をすべての参照専用データベース の間で回覧しながら、すべての参照専用データベースに 更新情報を転送するので、更新情報を通知しなければな らないデータベースの中に障害による停止あるいはネッ 50

トワーク上の問題等により、通信ができないものを含ん でいる場合があっても、回覧リストの中から通信可能状 態にある相手を動的に選択できるので、通信可能なもの から順に速やかに更新情報を通知することができる。同 時に、更新情報の通知という作業を参照専用データベー スで共同して行うため、主複製を持つデータベースへの 負荷の集中が避けられる。

【0042】また、更新情報の緊急性に応じて緊急度の 高いもの程細かく更新情報の回覧リストを分割し、同時 に複数のデータベース間で更新情報を回覧するので、緊 急度の高い更新情報を迅速に通知することができる。例 えば、クレジットカードの住所変更情報と盗難情報では 更新の緊急度が異なるが、本発明では、このような更新 情報の緊急度に応じて回覧リストを細かく分割し、同時 に複数のデータベースの間で回覧することによって、緊 急度の高い更新情報をすべてのデータベースに迅速に伝 えることができる。

【0043】また、従来の同報を利用した更新情報の通 知方法では、同報を行ったときにネットワークの負荷が 急速に高まるため、ネットワークの負荷が安定しないと いう問題があったが、本発明では、ネットワークの負荷 状態を監視し、負荷が低いときにのみ回覧リストを細か く分割し、同時に複数のデータベースの間で回覧するこ とによって、ネットワークの利用の最適化を図ることが でき、通信ネットワークの最大負荷を低減することがで きる。

#### 【図面の簡単な説明】

【図1】本発明の一実施例に係わる分散データベース制 御方法が適用される分散データベースの構成を示すブロ ック図である。

【図2】図1の分散データベースに使用されている主複 製を有するデータベースの機能を示す論理ブロック図で ある。

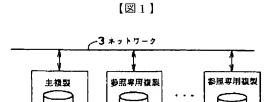
【図3】更新情報のデータ構造を示す図である。

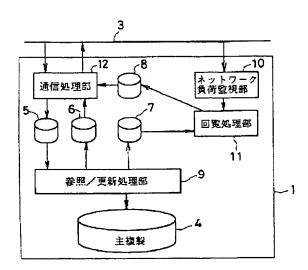
【図4】図1の分散データベースに使用されている参照 専用の複製を有するデータベースの機能を示す論理ブロ ック図である。

【図5】従来の更新情報を複製に通知する方法を示す説 明図である。

#### 【符号の説明】

- 1 主複製を有するデータベース
- 2 参照専用複製を有するデータベース
- 3 通信ネットワーク
- 4 主複製
- 9 参照/更新処理部
- 10,25 ネットワーク負荷監視部
- 11,26 回覧処理部
- 18 参照専用の複製
- 2 3 参照処理部
- 24 更新処理部





【図2】

【図3】

